Faster Annealing Schedules for Quantum Annealing

Satoshi Morita

Department of Physics, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo 152-8551

New annealing schedules for quantum annealing are proposed based on the adiabatic the-
orem. These schedules exhibit faster decrease of the excitation probability than a linear
schedule. To derive this conclusion, the asymptotic form of the excitation probability for
quantum annealing is explicitly obtained in the limit of long annealing time. Its first-order
term, which is inversely proportional to the square of the annealing time, is shown to be de-
termined only by the information at the initial and final times. Our annealing schedules make
it possible to drop this term, thus leading to a higher order (smaller) excitation probability.
We verify these results by solving numerically the time-dependent Schrödinger equation for
small size systems

KEYWORDS: quantum annealing, annealing schedule, adiabatic theorem, quantum algorithm,
optimization problem

1. Introduction

Quantum annealing (QA) is a novel generic algorithm to solve optimization problems.[1–3]
As can be seen from the fact that QA is also called quantum adiabatic evolution in the field of
quantum computation,[4] QA is based on the adiabatic theorem of quantum mechanics.[5] One
considers the time-dependent Hamiltonian consisting of two terms, the potential term and
kinetic energy term, the latter being introduced by hand to solve the minimization problem of
the former. The Hamiltonian initially has only the kinetic energy term and smoothly changes
to eventually consist only of the potential term at the end of a process. If the Hamiltonian
varies slowly, the adiabatic theorem guarantees that the system follows the ground state of the
instantaneous Hamiltonian during the process. The idea of QA is that the non-trivial target,
the ground state of the potential energy, is obtained after the adiabatic evolution starting
from the trivial ground state of the kinetic energy.

An important question is how the error rate of QA depends on the annealing time τ. If
τ is too short, the change of the Hamiltonian is rapid and then the non-adiabatic transition
may occur. Consequently the algorithm would miss the desired solution. An earlier discussion
suggested that the residual energy defined by the energy difference between the solution
obtained by an algorithm and true one is proportional to $(\log \tau)^{-\zeta}$ with $\zeta \sim 6$.[6] While this
result is still under debate,[7, 8] the asymptotic behavior of the excitation probability in the large
τ limit is well-established.[8] According to the adiabatic theorem,[5] the excitation probability
after the adiabatic evolution obeys the power law, τ^{-2}.

To implement QA and derive the τ^{-2} error rate, one ordinarily assumes that the Hamiltonian depends linearly on time. However, there is no theoretical restriction on the annealing schedule, the changing rate of the Hamiltonian. In the present paper, we derive the explicit expression of the coefficient of the τ^{-2} term in the excitation probability without this restriction of linearity. It is surprising that its upper bound is determined only by the information at the initial and final times. Based on this result, we find the condition that this coefficient vanishes and propose new faster annealing schedules having a much faster rate τ^{-2m} of error decrease.

This article is organized as follows. In the next section, the upper bound of the excitation probability is derived for a general quantum system. We apply in §3 this result to QA and propose faster annealing schedules. In §4, we verify them by numerical simulations for the two-level system, the transverse field Ising model and the database search problem. The last section is devoted to conclusion.

2. Upper bound for excitation probability

Let us consider the general Hamiltonian which depends on time t only through the dimensionless time $s = t/\tau$,

$$H(t) \equiv \tilde{H}(t/\tau) = \tilde{H}(s), \tag{1}$$

where τ is the total annealing time. The state vector follows the Schrödinger equation,

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle, \tag{2}$$

or, in terms of the dimensionless time,

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}s}|\tilde{\psi}(s)\rangle = \tau\tilde{H}(s)|\tilde{\psi}(s)\rangle, \tag{3}$$

where we set $\hbar = 1$. We assume that the initial state is chosen to be the ground state of $\tilde{H}(0)$ and that the ground state of $\tilde{H}(s)$ is not degenerate for $0 \leq s \leq 1$. Thus the state vector is expected to keep track of the instantaneous ground state of $\tilde{H}(s)$ for sufficiently large τ. The adiabatic theorem[5] provides a condition for the adiabatic evolution as

$$\tau \gg \max_{s,j} \{A_j(s)\}, \tag{4}$$

$$A_j(s) \equiv \frac{1}{\Delta_j(s)^2} \left| \langle j(s)| \frac{\mathrm{d}\tilde{H}(s)}{\mathrm{d}s} |0(s)\rangle \right|, \tag{5}$$

where $|j(s)\rangle$ denotes the jth instantaneous eigenstate of $\tilde{H}(s)$ with the eigenvalue $\varepsilon_j(s)$ and $\Delta_j(s) \equiv \varepsilon_j(s) - \varepsilon_0(s)$. We assume that $|0(s)\rangle$ is the ground state of $\tilde{H}(s)$.

Under the above adiabatic condition (4), the asymptotic expansion with respect to τ yields the excitation amplitude at $s = 1$ as

$$\langle j(1)|\tilde{\psi}(1)\rangle \simeq \int_0^1 \mathrm{d}s \frac{\mathrm{e}^{\mathrm{i}\tau \int_0^s \mathrm{d}s'\, \Delta_j(s')}}{\Delta_j(s)} \langle j(s)| \frac{\mathrm{d}\tilde{H}(s)}{\mathrm{d}s} |0(s)\rangle + O(\tau^{-2}). \tag{6}$$

Using the integration by parts, we obtain the upper bound for the excitation probability as

$$\left|\langle j(1)|\tilde{\psi}(1)\rangle\right|^2 \lesssim \frac{1}{\tau^2}\left\{A_j(0) + A_j(1)\right\}^2 + O(\tau^{-3}). \tag{7}$$

This formula implies that the coefficient of the τ^{-2} term is determined only by the state of the system at $s = 0$ and 1 and vanishes if $A_j(s)$ is zero at $s = 0$ and 1.

When the τ^{-2} term vanishes, a similar calculation yields the next order term of the excitation probability. If $\tilde{H}'(0) = \tilde{H}'(1) = 0$, we obtain

$$\left|\langle j(1)|\tilde{\psi}(1)\rangle\right|^2 \lesssim \frac{1}{\tau^4}\left\{A_j^{(2)}(0) + A_j^{(2)}(1)\right\}^2 + O(\tau^{-5}), \tag{8}$$

where we defined

$$A_j^{(m)}(s) \equiv \frac{1}{\Delta_j(s)^{m+1}}\left|\langle j(s)|\frac{\mathrm{d}^m \tilde{H}(s)}{\mathrm{d}s^m}|0(s)\rangle\right|. \tag{9}$$

It is easy to see that the τ^{-4} term also vanishes when $\tilde{H}''(0) = \tilde{H}''(1) = 0$. We generalize these results as follows: If the kth derivative of $\tilde{H}(s)$ is equal to zero at $s = 0$ and 1 for all $k = 1, 2, \cdots, m-1$, the excitation probability has the upper bound

$$\left|\langle j(1)|\tilde{\psi}(1)\rangle\right|^2 \lesssim \frac{1}{\tau^{2m}}\left\{A_j^{(m)}(0) + A_j^{(m)}(1)\right\}^2 + O(\tau^{-2m-1}), \tag{10}$$

which is proportional to τ^{-2m}.

3. Faster annealing schedules

Although we have so far considered the general time-dependent Hamiltonian, the ordinary Hamiltonian for QA is composed of the potential term and the kinetic energy term,

$$\tilde{H}(s) = \left\{1 - f(s)\right\}H_{\mathrm{kin}} + f(s)H_{\mathrm{pot}}. \tag{11}$$

The function $f(s)$, representing the annealing schedule, satisfies $f(0) = 0$ and $f(1) = 1$. Thus $\tilde{H}(0) = H_{\mathrm{kin}}$ and $\tilde{H}(1) = H_{\mathrm{pot}}$. The ground state of H_{pot} corresponds to the optimal solution for the optimization problem. The kinetic energy is chosen so that its ground state is trivial. Therefore the above Hamiltonian connects the trivial initial state and the non-trivial desired solution.

The condition for the τ^{-2m} term to exist is obtained straightforwardly from the results of the previous section because the Hamiltonian (11) depends on time only through the annealing schedule $f(s)$. It is sufficient that the kth derivative of $f(s)$ is zero at $s = 0$ and 1 for $k = 1, 2, \cdots, m-1$. We note that $f(s)$ should belong to C^m, that is, $f(s)$ is an mth differentiable function whose mth derivative is continuous.

Examples of the annealing schedules $f_m(s)$ with the τ^{-2m} error rate are the following polynomials:

$$f_1(s) = s, \tag{12}$$

$$f_2(s) = s^2(3 - 2s), \tag{13}$$

$$f_3(s) = s^3(10 - 15s + 6s^2), \tag{14}$$

$$f_4(s) = s^4(35 - 84s + 70s^2 - 20s^3). \tag{15}$$

The linear annealing schedule $f_1(s)$, which shows the τ^{-2} error, has ordinarily been used in the past studies.[2,3] Although we here list only polynomials symmetrical with respect to the point $s = 1/2$, this is not essential. For example, $f(s) = (1 - \cos(\pi s^2))/2$ also has the τ^{-4} error rate because $f'(0) = f'(1) = f''(0) = 0$ but $f''(1) = -2\pi^2$.

4. Numerical results

To confirm the upper bound for the excitation probability discussed in the previous section, it is instructive to study the two-level system with the Hamiltonian,

$$H_{\mathrm{LZ}}(t) = -\left(\frac{1}{2} - f\left(\frac{t}{\tau}\right)\right)h\sigma^z - \alpha\sigma^x, \tag{16}$$

where the σ^α $(\alpha = x, y, z)$ are the Pauli matrices. The energy gap of H_{LZ} has the minimum 2α at $f(s) = 1/2$. If the annealing time τ is not large enough to satisfy eq. (4), the non-adiabatic transition occurs. The Landau-Zener theorem[9,10] provides the excitation probability $P_{\mathrm{ex}} = \left|\langle 1(1)|\tilde{\psi}(1)\rangle\right|^2$ as

$$P_{\mathrm{ex}} = \exp\left(-\frac{\pi\alpha^2\tau}{f'(s^*)h}\right), \tag{17}$$

where s^* denotes a solution for $f(s^*) = 1/2$. On the other hand, if τ is sufficiently large, the system evolves adiabatically during the process. Thus the excitation probability has the upper bound (10), which is estimated as

$$P_{\mathrm{ex}} \lesssim \frac{4h^2\alpha^2}{\tau^{2m}(h^2 + 4\alpha^2)^{m+2}}\left\{\left|\frac{\mathrm{d}^m f}{\mathrm{d}s^m}(0)\right| + \left|\frac{\mathrm{d}^m f}{\mathrm{d}s^m}(1)\right|\right\}^2. \tag{18}$$

We numerically solved the Schrödinger equation (2) for this system (16) with the Runge-Kutta method. Figure 1 shows the result for the excitation probability with annealing schedules (12)-(15). The initial state is the ground state of $H_{\mathrm{LZ}}(0)$. The parameters are chosen to be $h = 2$ and $\alpha = 0.2$. The curved and straight lines show eqs. (17) and (18), respectively. In the small and large τ regions, the excitation probability perfectly fits those two expressions.

We carried out simulations of a rather large system, the Ising spin system with random interactions. The quantum fluctuations are introduced by the uniform transverse field. Thus, the potential and kinetic energy terms are defined by

$$H_{\mathrm{pot}} = -\sum_{\langle ij\rangle} J_{ij}\sigma_i^z\sigma_j^z - h\sum_{i=1}^N \sigma_i^z, \tag{19}$$

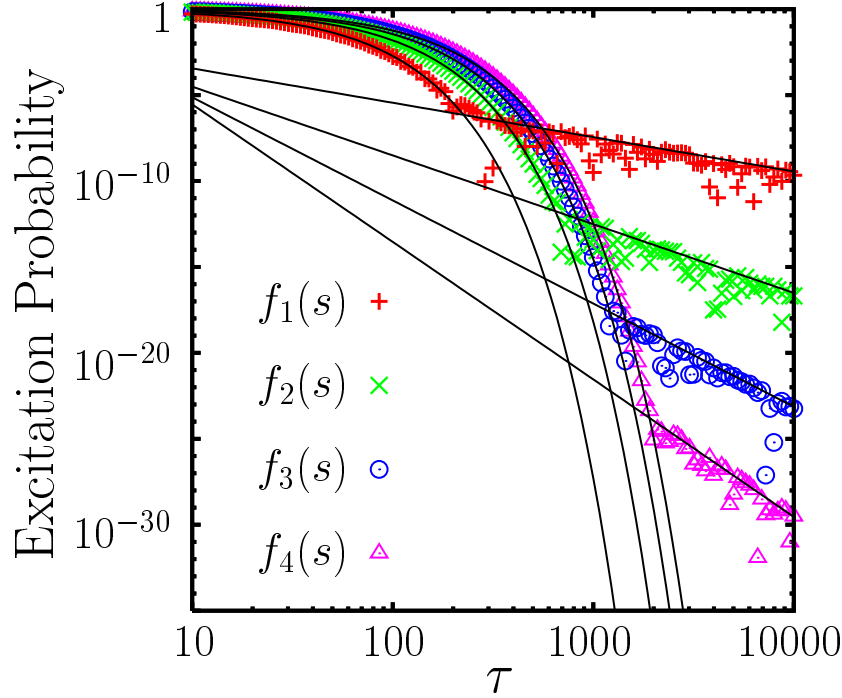$$H_{\mathrm{kin}} = -\Gamma\sum_{i=1}^N \sigma_i^x. \tag{20}$$

Fig. 1. The annealing-time dependence of the excitation probability for the two-level system (16) using schedules (12) to (15). The curved and straight lines show eqs. (17) and (18) for each annealing schedule, respectively. The parameters in eq. (16) are chosen to be $h = 2$ and $\alpha = 0.2$.

The initial state, namely the ground state of H_{kin}, is the all-up state along the x axis,

$$|\psi(0)\rangle = \bigotimes_i \frac{|z+\rangle_i + |z-\rangle_i}{\sqrt{2}}, \tag{21}$$

where $|z+\rangle_i$ and $|z-\rangle_i$ stand for the eigenstates of σ_i^z with eigenvalues 1 and -1, respectively.

The residual energy E_{res}, the energy difference between the solution obtained by the QA process and the exact one, is a useful measure for the error rate of QA. It has the same behavior as the excitation probability because it is rewritten as

$$E_{\mathrm{res}} \equiv \langle \tilde{\psi}(1)|H_{\mathrm{pot}}|\tilde{\psi}(1)\rangle - \varepsilon_0(1) \tag{22}$$

$$= \sum_{j>0} \Delta_j(1) \left| \langle j(1)|\tilde{\psi}(1)\rangle \right|^2. \tag{23}$$

Therefore E_{res} is expected to be asymptotically in proportion to τ^{-2m} using faster annealing schedules.

We investigated the two-dimensional square lattice of size 3×3. The quenched random coupling constants $\{J_{ij}\}$ are chosen from the uniform distribution between -1 and $+1$, as shown in Fig. 2. The parameters are $h = 0.1$ and $\Gamma = 1$. Figure 3 shows the τ dependence of the residual energy using annealing schedules (12)-(15). Straight lines representing τ^{-2m} ($m = 1, 2, 3, 4$) are also shown for comparison. The data clearly indicates the τ^{-2m}-law for
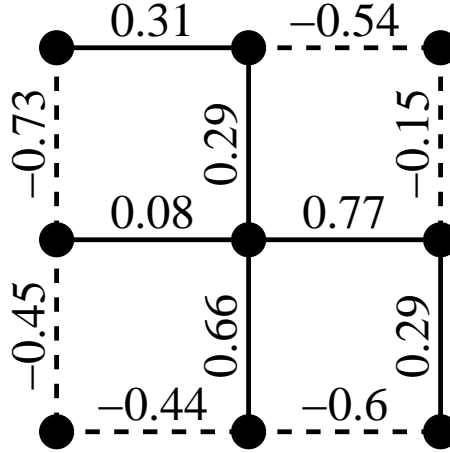
Fig. 2. Configuration of random interactions $\{J_{ij}\}$ on the 3×3 square lattice which we investigated. The solid and dashed lines indicate ferromagnetic and antiferromagnetic interactions, respectively.

large τ. Note that the irregular behavior around $E_{\text{res}} \sim 10^{-25}$ comes from numerical rounding errors.

Next, we apply the faster annealing schedule to the database search problem, finding an item in an unsorted database. Consider N items, among which one is marked. The goal of this problem is to find the marked item in a minimum time. The pioneering quantum algorithm proposed by Grover[11] solves this task in time of order $\sqrt{N}$, whereas the classical algorithm tests $N/2$ items on average. Farhi *et al.*[4] proposed a QA version of Grover's algorithm and Roland and Cerf[12] found an annealing schedule which has the same complexity as Grover's algorithm. Although their schedule is optimal in the sense that the excitation probability by the adiabatic transition is equal to a small constant at each time, it has the τ^{-2} error rate. We show that annealing schedules with the τ^{-2m} error rate can be constructed by a slight modification of their optimal schedule.

Let us consider the Hilbert space which has the basis states $|i\rangle$ $(i = 1, 2, \cdots, N)$, and the marked state is denoted by $|m\rangle$. Suppose that we can construct the Hamiltonian (11) with two energy terms,

$$H_{\text{pot}} = 1 - |m\rangle\langle m|, \tag{24}$$

$$H_{\text{kin}} = 1 - \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} |i\rangle\langle j|. \tag{25}$$

The Hamiltonian H_{pot} can be applied without the explicit knowledge of $|m\rangle$, which is the same assumption as in Grover's algorithm. The initial state is a superposition of all basis states,

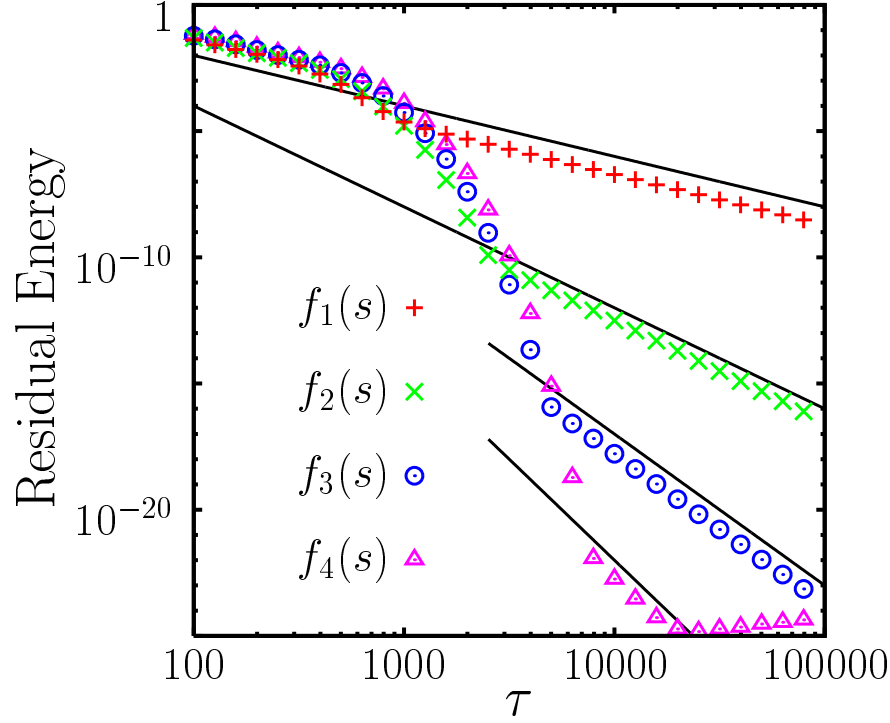$$|\psi(0)\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |i\rangle, \tag{26}$$

Fig. 3.　The annealing-time dependence of the residual energy for the two-dimensional spin glass model with faster annealing schedules. The solid lines stand for functions proportional to τ^{-2m} $(m = 1, 2, 3, 4)$. The parameters are $h = 0.1$ and $\Gamma = 1$.

which does not depend on the marked state. The energy gap between the ground state and the first excited state,

$$\Delta_1(s) = \sqrt{1 - 4\frac{N-1}{N}f(s)(1 - f(s))}, \tag{27}$$

has a minimum at $f(s) = 1/2$. The highest eigenvalue $\varepsilon_2(s) = 1$ is $(N - 2)$-fold degenerate.

To derive the optimal annealing schedule, we briefly review the results reported by Roland and Cerf.[12] When the energy gap is small, a non-adiabatic transition is likely to occur. Thus we need to change the Hamiltonian carefully. On the other hand, when the energy gap is not very small, too slow a change wastes time. Therefore the annealing schedule should be tuned to satisfy the adiabatic condition (4) in each infinitesimal time interval, that is,

$$\frac{A_1(s)}{\tau} = \delta, \tag{28}$$

where δ is a small constant. In the database search problem, this condition is rewritten as

$$\frac{\sqrt{N-1}}{\tau N \Delta_1(s)^3}\frac{\mathrm{d}f}{\mathrm{d}s} = \delta. \tag{29}$$

After integration under boundary conditions $f(0) = 0$ and $f(1) = 1$, we obtain that

$$f_{\mathrm{opt}}(s) = \frac{1}{2} + \frac{2s - 1}{2\sqrt{N - (N-1)(2s-1)^2}}. \tag{30}$$
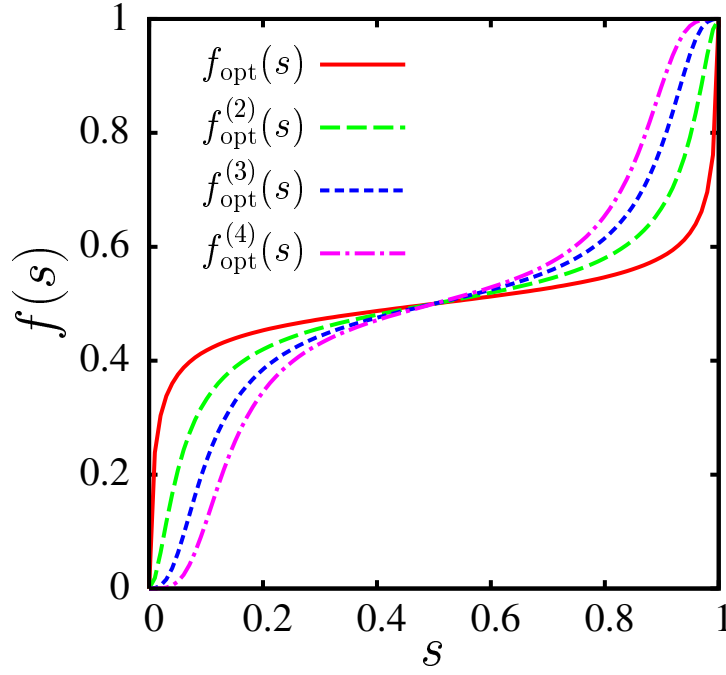
Fig. 4.   The optimal annealing schedules for the database search problem ($N = 64$). The solid line denotes the original optimal schedule (30) and the dashed lines stand for modified schedules.

As plotted by a solid line in Fig. 4, this function changes most slowly when the energy gap takes a minimum value. It is noted that the annealing time is determined by the small constant δ as

$$\tau = \frac{\sqrt{N - 1}}{\delta}, \tag{31}$$

which means that the computation time is of order $\sqrt{N}$ similarly to Grover's algorithm.

The optimal annealing schedule (30) shows the τ^{-2} error rate because its derivative is not zero at $s = 0$ and 1. It is easy to see from eq. (30) that a simple replacement of s with $f_m(s)$ fulfills the condition for the τ^{-2m} error rate. We carried out numerical simulations for $N = 64$ with such annealing schedules, $f_{\text{opt}}^{(m)}(s) \equiv f_{\text{opt}}(f_m(s))$, as plotted by dashed lines in Fig. 4. As shown in Fig. 5, the residual energy with $f_{\text{opt}}^{(m)}(s)$ is proportional to τ^{-2m}. The characteristic time τ_c for the τ^{-2m} error rate to show up increases with m: Since the modified optimal schedule $f_{\text{opt}}^{(m)}(s)$ has a steeper slope at $s = 1/2$ than $f_{\text{opt}}(s)$, a longer annealing time is necessary to satisfy the adiabatic condition (29). Nevertheless, the difference in slopes of $f_{\text{opt}}^{(m)}(s)$ is only a factor of $O(1)$, and therefore τ_c is still scaled as $\sqrt{N}$.

5.   Conclusion

In this paper, we have studied the error rate of QA based on the adiabatic theorem and proposed annealing schedules which present a faster decrease of the error rate with the anneal-
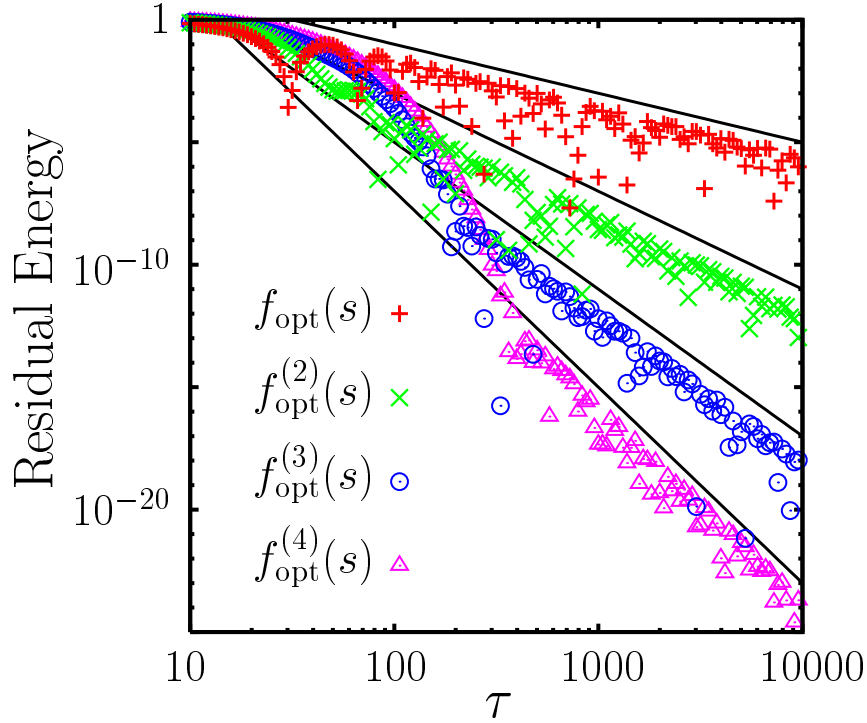
Fig. 5. The annealing-time dependence of the residual energy for the database search problem ($N = 64$) with the optimal annealing schedules plotted in Fig. 4. The solid lines stand for functions proportional to τ^{-2m} ($m = 1, 2, 3, 4$).

ing time. The point is that the excitation probability is determined only by the information at $t = 0$ and τ when the system evolves adiabatically from the beginning to the end. Thus one can easily remove the usual τ^{-2} error by tuning the annealing schedule. As examples of such schedules, we listed a few polynomials.

Although our new faster schedules do not improve the characteristic time τ_c for the adiabatic evolution, these schedules can be used to produce a new faster schedule from the known efficient schedule with small τ_c. In general, it is difficult to obtain the optimal annealing schedule analytically. Nevertheless, if one once finds an efficient schedule for a specific problem, schedules with smaller error rates are immediately constructed by a slight modification under the present scheme.

Acknowledgements

References

1) T. Kadowaki and H. Nishimori: Phys. Rev. E **58** (1998) 5355.

2) A. Das and B. K. Chakrabarti: *Quantum Annealing and Related Optimization Methods* (Springer, Berlin Heidelberg, 2005) Lecture Notes in Physics, Vol. 679.

3) G. E. Santoro and E. Tossati: J. Phys. A: Math. Gen. **39** (2006) R393.

4) E. Farhi, J. Goldstone, S. Gutmann and M. Sipser: quant-ph/0001106.

5) A. Messiah: Quantum Mechanics (Wiley, New York, 1976).

6) G. E. Santoro, R. Martoňák, E. Tosatti and R. Car: Science **295** (2002) 2427.

7) M. Sarjala, V. Petäjä and M. Alava: J. Stat. Mech. (2006) P01008.

8) S. Suzuki and M. Okada: J. Phys. Soc. Jpn **74** (2005) 1649.

9) L. D. Landau and E. M. Lifshitz: *Quantum Mechanics: non-relativistic theory*, 2nd ed. (Pergamon Press, Oxford, 1965).

10) C. Zener: Proc. R. Soc. London, Ser. A **137** (1932) 696.

11) L. K. Grover: Phys. Rev. Lett. **79** (1997) 325.

12) J. Roland and N. J. Cerf: Phys. Rev. A **65** (2002) 042308.